197

Reihe Ökonomie

Economics Series

The Carbon Kuznets Curve:

A Cloudy Picture Emitted by Bad Econometrics?

Martin Wagner

197

Reihe Ökonomie

Economics Series

The Carbon Kuznets Curve:

A Cloudy Picture Emitted by Bad Econometrics?

Martin Wagner

November 2006

Institut für Höhere Studien (IHS), Wien
Institute for Advanced Studies, Vienna

Contact:

Martin Wagner
Department of Economics and Finance
Institute for Advanced Studies
Stumpergasse 56
1060 Vienna, Austria
☎: +43/1/599 91-150
fax: +43/1/599 91-163
email: martin.wagner@ihs.ac.at

Abstract

In recent years many empirical studies of environmental Kuznets curves employing unit root and cointegration techniques have been conducted for both time series and panel data. When using such methods several issues arise: the effects of a short time dimension, in a panel context the effects of cross-sectional dependence, and the presence of nonlinear transformations of integrated variables. We discuss and illustrate how ignoring these problems and applying standard methods leads to questionable results. Using an estimation approach that addresses the second and third problem we find no evidence for an inverse U-shaped relationship between GDP and CO_2 emissions.

Keywords

Carbon Kuznets Curve, panel data, unit roots, cointegration, crosssectional dependence, nonlinear transformations of regressors

JEL Classification

C12, C13, Q20

Comments

The comments of Klaus Neusser, Georg Müller-Fürstenberger, and Reto Tanner are gratefully acknowledged. The usual disclaimer applies.

Contents

1 Introduction 1

2 The Carbon Kuznets Curve 5

3 Panel Unit Root Tests 7

3.1 First Generation Tests 8

3.2 Second Generation Tests 12

3.3 Conclusions from Panel Unit Root Analysis 16

4 Panel Cointegration Tests 17

5 Estimation of the Carbon Kuznets Curve with Panel Cointegration Methods and Using De-factored Observations 21

5.1 Panel Cointegration Estimation 21

5.2 Estimation with De-Factored Observations 23

6 Summary and Conclusions 24

References 26

Appendix: Data and Sources 32

Appendix B: Bootstrap Algorithms 33

1 Introduction

Besides nuclear energy, hydrocarbon deposits like petroleum, coal and natural gas are currently the only available large scale primary energy sources. Their utilization as fossil fuels leads to the emission of – amongst other pollutants – CO_2, which is considered the principal anthropogenic greenhouse gas. Since most economic activities require the use of energy, a link between economic activity and CO_2 emissions appears plausible.

Increased atmospheric CO_2 concentration can persist up to thousands of years. It exerts a warming influence on the lower atmosphere and the surface, i.e. it initiates climate change, see Peixoto and Ort (1992) or Ramanathan, Cicerone, Singh, and Kiehl (1985). Rational and efficient climate policy requires reliable understanding and accurate quantification of the link between economic activity and CO_2 emissions.

In this paper we are concerned with the econometric analysis of the relationship between GDP and emissions. The core of the econometric approach to study the link between GDP and CO_2 emissions usually consists of estimating a reduced form relationship on cross-section, time series or panel data sets. Estimation techniques as well as variables chosen vary substantially across studies. Most of the studies focus on a specific conjecture, the so-called 'Environmental Kuznets Curve' (EKC) hypothesis. This hypothesis claims an inverse U–shaped relation between (the logarithm of per capita) GDP and pollutants. In the specific case of CO_2 emissions we speak of the 'Carbon Kuznets Curve' (CKC).[1]

The EKC hypothesis has been initiated by the seminal work of Gene Grossman and Alan Krueger (1991, 1993, 1995). They postulate, estimate and ascertain an inverse U–shaped relationship between measures of several pollutants and per capita GDP.[2] Summary discussions of this empirical literature are contained in Stern (2004) or Yandle, Bjattarai, and Vijayaraghavan (2004), who find more than 100 refereed publications of this type.[3]

[1] Note that also specifications in levels instead of logarithms are used in the literature.

[2] To be precise, Grossman and Krueger actually use a third order polynomial in GDP whereas the quadratic specification seems to have been initiated by Holtz-Eakin and Selden (1995).

[3] A prominent alternative approach to study the links between economic activity and environmental damages in general or emissions in particular is given by 'Integrated Assessment Models', pioneered with DICE of Nordhaus (1992) or MERGE by Manne, Mendelsohn, and Richels (1995). This approach consists of specifying and calibrating a general equilibrium model of the world economy. The economic model is then linked with a climate model to integrate the effects of climate change feedbacks into the economic analysis. To a certain extent the econometric and the integrated assessment model approach can be seen as complements. Unfortunately, only few authors have tried to combine the two approaches, see McKibbin, Ross, Shackleton, and Wilcoxen (1999) for one example. Müller-Fürstenberger and Wagner (2006) contains a discussion on the relation or lack thereof between reduced form econometric findings and relationships derived with structural models.

In the empirical EKC literature there is an ongoing discussion on appropriate specification and estimation strategies, see Dijkgraaf and Vollebergh (2005) for a comparative discussion of econometric techniques applied in the literature. It is the aim of this study to contribute to this discussion by addressing several serious econometric problems that have not been appropriately handled or have been ignored to a certain extent up to now. We focus on parametric approaches only. For non-parametric EKC approaches (see e.g. Millimet, List, and Stengos, 2003), semi-parametric approaches (see e.g. Bertinelli and Strobl, 2005) or versions based on spline interpolation (see e.g. Schmalensee, Stoker, and Judson, 1998). To illustrate our arguments, we present computations for a panel data set for the Carbon Kuznets Curve comprising 107 countries (see Table 7 in Appendix A) over the period 1986–1998.

The discussion is on two – related – levels. The *first level* is a fundamental discussion on whether the time series and panel EKC literature is applying the appropriate tools. The *second level* is the issue whether the tools applied – abstracting from the first level issue of appropriateness – are applied correctly or with enough care. Of course, those two issues are related and there will be substantial overlap in the two levels of discussion. We turn to both issues below, but can already present the main observation here: The answer is rather negative on both levels.

When using time series or panel data the issue of stationarity of the variables is of prime importance for econometric analysis. This is due to the fact that the properties of many statistical procedures depend crucially upon stationarity or unit root nonstationarity, i.e. integratedness, of the variables used. Related to this issue is the question of spurious regression (see e.g. Phillips, 1986) versus cointegration, see the discussion below. One part of the literature, in particular the early literature, completely ignores this issue, see e.g. Grossmann and Krueger (1991), Grossmann and Krueger (1995), Holtz-Eakin and Selden (1995) or Martinez-Zarzoso and Bengochea-Morancho (2004) to name just a few.[4]

Another part of the literature is mentioning the stationarity versus unit root nonstationarity issue, these include inter alia Perman and Stern (2003), Stern (2004); and when allowing also for breaks Heil and Selden (1999) or Lanne and Liski (2004) (the latter in a time series context) are two examples. The problem is, however, that three important issues – on both levels of our discussion - have been ignored thus far. On the first level these are the following

[4]Two further empirical issues are neglected in this paper, since they are in principle well understood. These are the homogeneity of the relationship for large heterogeneous panels and the question of stability of estimated relationships.

two – given that the variables are indeed unit root nonstationary. First, the usual formulation of the EKC involves squares or even third powers of (log) per capita GDP. If (log) per capita GDP is integrated, then nonlinear transformations of it, as well as regressions involving such transformed variables, necessitate a different type of asymptotic theory and also lead to different properties of estimators. Regression theory with nonlinear transformations of integrated variables has only recently been studied in Chang, Park, and Phillips (2001), Park and Phillips (1999) and Park and Phillips (2001). Currently no extension of these methods to the panel case is available, which posits a fundamental challenge to the empirical EKC literature.[5] To our knowledge this nonlinearity issue has not been discussed at all in the EKC literature. One study avoiding the above problems is given by Bradford, Fender, Shore, and Wagner (2005). These authors base their results, using the Grossman and Krueger (1995) data, on an alternative specification comprising instead of income over time only an average level and the average growth rate of income. Thus, this study circumvents the problems arising in regressions containing nonlinear transformations of nonstationary regressors.

Second, in case of nonstationary panel analysis, all the methods used so far in the EKC literature rely upon the *cross-sectional independence* assumption. I.e. these, so called 'first-generation' methods assume that the individual countries' GDP and emissions series are independent across countries. This rather implausible assumption is required for the first generation methods to allow for applicability of simple limit arguments (along the cross-section dimension). In this respect progress has been made in the theoretical literature and several panel unit root tests that allow for cross-sectional dependence are available. Several such tests are applied in this study, which seems to be the first application of such 'second-generation' methods in the EKC context.

Third, on the second level of discussion the major issue is the following: The 'first-generation' methods used for nonstationary panels are known to perform very poor for short panels. This stems from the fact that the properties of the panel unit root and cointegration tests crucially depend on the properties of the methods used at the individual country level. If the panel method is based on pooling, then the very poor properties of time series unit root tests for short time series feed directly into bad properties of pooled panel unit root tests, see

[5]To be precise: We do not claim that e.g. estimation of a quadratic CKC with integrated regressors by some panel cointegration estimator is inconsistent. We just want to highlight that the (linear cointegration) methods are not designed for such problems and that nonlinear transformations of integrated variables have fundamentally different asymptotic behavior than integrated properties. These two aspects imply that it is up to now unclear what such results could mean, or which properties such results have.

Hlouskova and Wagner (2006a) for ample simulation evidence. We show in this paper that by applying bootstrap methods – ignoring as mentioned above the more fundamental question of applicability of such first-generation methods at that point – quite different results than based on asymptotic critical values can be obtained. We have implemented three different bootstrap algorithms that are briefly described in Appendix B. These are the so called parametric, the non-parametric and the residual based block (RBB) bootstrap. The RBB bootstrap has been developed for non-stationary time series by Paparoditis and Politis (2003). The first two methods obtain white noise bootstrap replications of residuals due to pre-whitening and the latter is based on re-sampling blocks of residuals to preserve the serial correlation structure. The difference between the parametric and the non-parametric bootstrap is essentially that in the former the residuals are drawn from a normal distribution while in the latter they are re-sampled from the residuals.

It seems that the uncritical use of asymptotic critical values might be a main problem at the second level of discussion we intend to initiate with this paper. Even stronger, we find that one can support any desired result concerning unit root and cointegration behavior by choosing the test (and to a certain extent the bootstrap algorithm) 'strategically'. Furthermore and related to the above, standard panel cointegration estimation results of the CKC differ widely across methods. These findings cast serious doubt on the results reported so far in the literature – even when ignoring the two first level problems (nonlinear transformations, cross-sectional correlations). We include this type of discussion to show that, even when ignoring the first level problems and staying within the standard framework applied up to now, the empirical (panel and time series) EKC literature is an area where best econometric practice is generally not observed.

The paper is organized as follows: In Section 2 we briefly discuss the specification of the CKC and set the stage for the subsequent econometric analysis. In Section 3 we discuss first- and second-generation panel unit root test results, and in Section 4 we discuss panel cointegration test results. Section 5 presents the results of CKC estimates based on panel cointegration methods and based on de-factorized data. Section 6 briefly summarizes and concludes. Two appendices follow the main text. In Appendix A we describe the data and their sources. Appendix B briefly describes the implemented bootstrap procedures.

2 The Carbon Kuznets Curve

In our parametric CKC specification we focus on the logarithms of both per capita GDP, denoted by y_{it}, and per capita CO_2 emissions, denoted by e_{it}.[6] Here and throughout the paper $i = 1, \ldots, N$ indicates the country and $t = 1, \ldots, T$ is the time index. Qualitatively similar results have also been obtained when using levels instead of logarithms.

Our sample encompasses 107 countries, listed in Table 7 in Appendix A, over the years 1986–1998. The major region omitted is the former Soviet Union and some other formerly centrally planned economies. We also exclude countries with implausibly huge jumps in emissions or GDP, as it is the case for Kuwait for example.[7]

The basic formulation of the CKC in logarithms we focus on, is given by

$$\ln(e_{it}) = \alpha_i + \gamma_i t + \beta_1 \ln(y_{it}) + \beta_2 (\ln(y_{it}))^2 + u_{it}, \tag{1}$$

with u_{it} denoting the stochastic error term, for which depending upon the test or estimation method applied different assumptions concerning serial correlation have to be made. In this formulation we include in general both fixed effects, α_i, and country specific linear trends, $\gamma_i t$. These linear trends are included to allow for exogenous decarbonization of GDP due to technical progress and structural change. We have also experimented with specifications that include time specific fixed effects, but these do not qualitatively change the results. Thus, we focus in this paper on specifications including fixed effects or fixed effects and trends, since these are the two common specifications of deterministic components in unit root and cointegration analysis. The above formulation of the CKC posits a strong *homogeneity* assumption. The functional form is assumed to be identical across countries, since the coefficients β_1 and β_2 are restricted to be identical across countries. Heterogeneity across countries is only allowed via the fixed effects and linear trends. Different α_i shift the overall level of the relationship, and different trend slopes γ_i across countries shift the quadratic relationship differently across countries over time. This, of course, might be too restrictive for a large panel with very heterogeneous countries. See e.g. Dijkgraaf and Vollebergh (2005) for a discussion (and rejection) of homogeneity for a panel of 24 OECD countries.

Equation (1) allows to discuss one major overlooked problem related with potential non-

[6] Throughout the paper we are usually only concerned with logarithms of per capita GDP and emissions and will not always mention that explicitly.

[7] The carbon data have been multiplied by 1000 to convert them into kilos, which results in data of the same order of magnitude as the GDP data measured in dollars.

stationarity of emissions and/or GDP, namely that of nonlinear transformations of integrated regressors. The macro-econometric literature has gathered a lot of evidence that in particular GDP series are very likely *integrated.* A stochastic process, x_t say, is called integrated, if its first difference, $\Delta x_t = x_t - x_{t-1}$ is stationary, but x_t is not. Let ε_t denote a *white noise* process. Then the simplest integrated process is given by the random walk, i.e. by accumulated white noise, $x_t = \sum_{j=1}^{t} \varepsilon_j$.[8] By construction the first difference of x_t is white noise. Now, what about the first difference of x_t^2? Straightforward computations give $\Delta x_t^2 = \Delta \left(\sum_{j=1}^{t} \varepsilon_j \right)^2$ equal to $\Delta x_t^2 = \varepsilon_t^2 + 2\varepsilon_t \sum_{j=1}^{t-1} \varepsilon_j$. Thus, as expected, the first difference of the square of an integrated process is not stationary. The relationship to the CKC is clear: Both the logarithm of per capita GDP and its square are contained as regressors. However, at most one of them can be an integrated process. This fact has been overlooked in the CKC literature up to now.[9]

The above problem is fundamental and no estimation techniques for panel regressions with nonlinear transformations of integrated processes are available. Only recently there has been a series of papers by Peter Phillips and coauthors that addresses this problem for time series observations. This literature shows that the asymptotic theory required, as well as they asymptotic properties obtained, generally differ fundamentally from the standard integrated case.[10] However, we nevertheless will present in the sequel unit root and cointegration tests with the quadratic specification as given in (1) to show that the cointegration techniques have probably not been applied with enough care. We perform bootstrap inference for unit root and cointegration tests to show that the asymptotic critical values are bad approximations to the finite sample critical values. Thus, we argue, that even when being unaware of the first level problems, a more critical application of standard techniques would lead a researcher in good faith to use the proper toolkit to be more cautious about the results.

As a benchmark case, where we avoid the issue of nonlinear transformations of integrated regressors, we also include the linear specification (2) in our analysis. It is only this linear case for which the panel unit root and cointegration tests can be applied with a sound theoretical

[8] Here and throughout we ignore issues related to starting values as they are inessential to our discussion.

[9] Several authors, e.g. Perman and Stern (2003), even present unit root test results on log per capita GDP and its square. Furthermore they even present 'cointegration' estimates of the EKC. This does not have a sound econometric basis. Consistent estimation techniques for this type of estimation problem have to be established first.

[10] Relevant papers are Park and Phillips (1999), Chang, Park, and Phillips (2001) and Park and Phillips (2001). Current research is concerned with an application of these theoretical results to the EKC/CKC hypothesis.

basis, given that log per capita GDP is indeed integrated.

$$\ln(e_{it}) = \alpha_i + \gamma_i t + \beta_1 \ln(y_{it}) + u_{it} \tag{2}$$

The second first level issue is that all the EKC papers that use panel unit root or cointegration techniques only apply so called 'first generation' methods. These methods require that the regressors and the errors in the individual equations are independent across countries. In this paper we present the first application of 'second generation' panel unit root tests that allow for cross-sectional dependence. Indeed strong evidence for cross-sectional dependence is found, discussed in Section 3.2. In the following sections, to parallel the historical development of methods, we nevertheless will start with reporting the results obtained by bootstrapping first generation methods. All results, and in particular the first generation results, have to be seen in the light of the critical issues this paper is concerned about.

3 Panel Unit Root Tests

The time dimension of the sample with only 13 years necessitates the application of panel unit root tests. The section is split in two subsections. In subsection 3.1 we discuss first generation tests that rely upon the assumption of cross-sectional independence. So far, only this type of test has been used in the EKC literature. In particular we show that a straightforward application of such tests can be misleading, since the finite sample distribution of the test statistics can differ substantially from the asymptotic distribution. This implies that inference based on the asymptotic critical values can be misleading, see Hlouskova and Wagner (2006a) for large scale simulation evidence in this respect. Panel unit root tests should therefore only be applied with great care.

In subsection 3.2 we report results obtained by applying second-generation panel unit root tests. We find strong evidence for cross-sectional correlation. Of course, these second generation methods should be applied first, and only when no cross-sectional correlation is found, one can resort to first generation methods. We revert this logical sequence to show that conditionally upon staying in the first generation framework, much more care than is common in the literature should be taken.

3.1 First Generation Tests

Let x_{it} denote the variable we want to test for a unit root, i.e. we want to test the null hypothesis $H_0 : \rho_i = 1$ for all $i = 1, \dots, N$ in

$$x_{it} = \rho_i x_{it-1} + \alpha_i + \gamma_i t + u_{it} \tag{3}$$

where u_{it} are stationary processes assumed to be cross-sectionally independent.[11] The tests applied differ with respect to the alternative hypothesis. The first alternative is the *homogenous* alternative $H_1^1 : \rho_i = \rho < 1$ (and bigger than -1) for $i = 1, \dots, N$. The *heterogeneous* alternative is given by $H_1^2 : \rho_i < 1$ for $i = 1, \dots, N_1$ and $\rho_i = 1$ for $i = N_1 + 1, \dots, N$.[12] Especially for heterogeneous panels the alternative H_1^2 might be the more relevant one. However, in the literature both alternatives have been used. In our data set we observe no systematic differences in the results between tests with the homogenous and the heterogeneous alternative, see the results below and in Table 1.

In general, some correction for serial correlation in u_{it} will be necessary. Two main approaches are followed in all tests, either a non-parametric correction in the spirit of Phillips and Perron (1988) or in the spirit of the augmented Dickey Fuller (ADF) principle. The ADF correction adds lagged differences of the variable (Δx_{it-j})to the regression to achieve serially uncorrelated errors.

The following tests have been implemented:[13] The test of Levin, Lin, and Chu (2002) (LL), which is after suitable first step corrections a pooled ADF test. The second is the test of Breitung (2000) (UB), which is a pooled ADF type test based on a simple bias correction. These two tests, due to their pooled estimation of ρ, test against the homogenous alternative. We have implemented three tests with the heterogeneous alternative. Two of them are developed by Im, Pesaran, and Shin (1997, 2003). One is given by essentially the group-mean of individual ADF t-statistics (IPS), and the other is a group-mean LM statistic ($IPS - LM$). Finally, we present one test based on the Fisher (1932) test principle. The idea of Fisher is to use the fact that under the null hypothesis the p-values of a continuous test statistic are uniformly distributed over the unit interval. Then, minus two times the logarithm of the p–values is distributed as χ^2_2. This implies that the sum of N independent

[11] Note that also time specific effects θ_t can be included.

[12] With $\lim_{N\to\infty} \frac{N_1}{N} > 0$.

[13] We abstain here from a discussion of the limit theory underlying the asymptotic results. Most of the results are based on sequential limit theory, where first $T \to \infty$ followed by $T \to \infty$.

transformed p–values is distributed as χ^2_{2N}.[14] We follow the work of Maddala and Wu (1999) (MW) and implement this idea by using the ADF test for each cross-sectional unit.

We furthermore report the Harris and Tzavalis (1999) test results. Their test is identical to the Levin, Lin, and Chu (2002) test, except for that Harris and Tzavalis derive the exact finite T test distribution. This may be advantageous for our short panel. The exact test distribution comes, however, at a high price. Harris and Tzavalis derive their results only for the case when u_{it} is white noise. All tests except for MW, which is χ^2_{2N} distributed, are asymptotically standard normally distributed. We perform tests with both the homogenous and the heterogeneous alternative to see whether there are big differences in the test behavior across these two tests. This, however, does not appear to be the case.

As mentioned already, it is known that for panels of the size available in this study (with T only equal to 13), the asymptotic distributions of panel unit root and panel cointegration tests provide poor approximations to the small sample distributions (see e.g. Hlouskova and Wagner, 2006a). Hence, the notorious size and power problems for which unit root tests are known in the time series context also appear in short panels. In Figure 1 we display the asymptotic null distribution (the standard normal distribution) and the bootstrap null distributions (from the non-parametric bootstrap) when testing for a unit root in CO_2 emissions including only fixed effects in the test specification, for the five asymptotically standard normally distributed tests. The figure shows *substantial* differences between the bootstrap approximations to the finite sample distribution of the tests and their asymptotic distribution. Thus, basing inference on the asymptotic critical values can lead to substantial size distortions. The discrepancy between the asymptotic and the bootstrap critical values can also be seen in Table 1, where the 5% bootstrap critical values are displayed in brackets. They vary substantially both across tests and also across the two variables. In most cases they are far away from the asymptotic critical values ± 1.645, respectively 249.128 for the Maddala and Wu test.

It is customary practice in unit root testing to test in specifications with and without linear trends included. A linear trend in the test equation, when there is no trend in the data generating process reduces the power of the tests. Conversely, omitting a trend when there is a trend in the data, induces a bias in the tests towards the null hypothesis. Graphical inspection

[14] By appropriate scaling and for $N \to \infty$, Choi (2001) derives asymptotically standard normally distributed tests based on this idea.

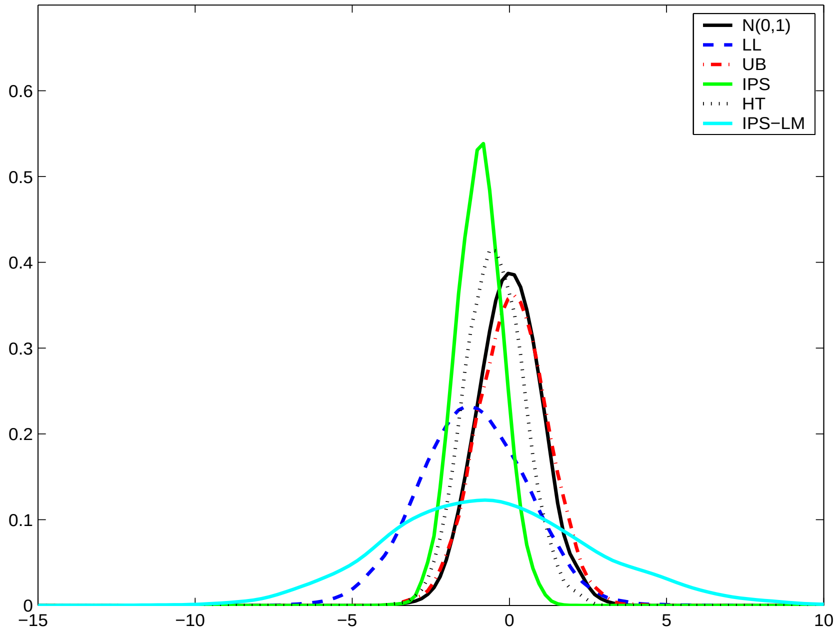

Figure 1: Bootstrap test statistic distributions for CO_2 for 5 asymptotically standard normally distributed panel unit root tests.
The results are based on the non-parametric bootstrap with 5000 replications. Fixed effects are included.

of the data leads us to conclude that for CO_2 emissions the specification without trend might be sufficient, whereas for GDP the specification with trend might be more appropriate. The nature of the trend component of GDP is a widely discussed topic in macro-econometrics. Both, unit root nonstationarity with its underlying stochastic trend or trend-stationarity with usually a linear deterministic trend are plausible and widely used specifications. This uncertainty concerning the trend specification for GDP manifests itself also in our panel test results, see below. For completeness we report both types of results for both variables. The first block in Table 1 displays the results for the parametric bootstrap, the second for the non-parametric bootstrap and the third for the RBB bootstrap. Within each of the blocks, the first block-row shows the results with fixed effects and the second the results when both fixed effects and linear trends are included.

PARAMETRIC BOOTSTRAP						
Variable	*LL*	*UB*	*IPS*	*HT*	*IPS − LM*	*MW*
	Fixed Effects					
CO_2	**-2.807*** (-3.957)	0.915 (-2.159)	0.229 (-1.707)	**-4.828*** (-5.705)	-1.291 (1.096)	**310.781*** (313.176)
GDP	**-5.890** (-3.197)	1.512 (-2.626)	-1.590 (-0.582)	4.321 (3.216)	0.070 (1.231)	**422.513** (329.209)
	Fixed Effects and Trends					
CO_2	**-8.493** (-6.501)	-0.565 (-1.121)	**-2.093** (-1.823)	-8.618 (-12.711)	0.259 (0.276)	**418.543** (362.505)
GDP	**-15.911** (-2.635)	2.072 (-1.167)	**-3.423** (-1.346)	12.302 (4.419)	0.456 (0.301)	**530.792** (378.350)
NON-PARAMETRIC BOOTSTRAP						
	LL	*UB*	*IPS*	*HT*	*IPS − LM*	*MW*
	Fixed Effects					
CO_2	**-2.807** (-2.023)	0.915 (-4.166)	0.230 (-1.628)	**-4.828** (-1.029)	-1.291 (-1.343)	**310.781** (309.904)
GDP	**-5.890** (-1.775)	1.512 (-0.974)	-1.590 (2.070)	**4.321** (5.241)	0.070 (0.361)	**422.515** (323.413)
	Fixed Effects and Trends					
CO_2	**-8.493*** (-10.289)	-0.565 (-1.226)	**-2.094*** (-2.485)	**-8.618*** (-12.774)	0.259 (0.182)	**418.543** (403.105)
GDP	**-15.911** (-8.711)	2.072 (-1.176)	**-3.423** (-1.789)	**12.302** (13.777)	0.456 (0.201)	**530.792** (409.514)
RESIDUAL BASED BLOCK BOOTSTRAP						
	LL	*UB*	*IPS*	*HT*	*IPS − LM*	*MW*
	Fixed Effects					
CO_2	**-2.807*** (-7.603)	0.915 (-5.999)	0.230 (-4.094)	**-4.828*** (-8.351)	-1.291 (3.006)	**310.781*** (364.274)
GDP	**-5.890*** (-9.082)	1.512 (-6.344)	-1.590 (-4.896)	4.321 (-6.901)	0.070 (3.846)	**422.513** (392.093)
	Fixed Effects and Trends					
CO_2	**-8.493*** (-23.999)	-0.565 (-1.222)	-2.094 (-6.096)	**-8.618** (-8.462)	0.259 (4.226)	**418.543*** (608.021)
GDP	**-15.911*** (-18.717)	2.072 (-2.120)	**-3.423*** (-8.631)	12.302 (-5.887)	0.456 (4.694)	**530.792*** (663.504)

Table 1: Results of first generation panel unit root tests for the logarithm of per capita CO_2 emissions and the logarithm of per capita GDP including only fixed effects in the upper block-rows and fixed effects and linear trends in the lower block-rows. The first part of the table corresponds to the parametric bootstrap, the second to the non-parametric bootstrap and the third to the residual based block bootstrap. In parentheses the 5 % critical values obtained by the three different bootstrap methods are displayed.
The asymptotic 5 % critical value is given by -1.645 for the first 4 tests, by 1.645 for IPS-LM and by 249.128 for MW.
Bold indicates rejection of the null hypothesis based on the bootstrap critical values and **bold*** indicates rejection based upon the asymptotic critical values but no rejection according to the bootstrap critical values.
The autoregressive lag lengths in both the autoregression based tests, in the parametric bootstrap and the non-parametric bootstrap are equal to 1. The block-length in the RBB bootstrap is equal to 2.

Let us start with (the logarithm of per capita) CO_2 emissions. For all three bootstrap methods and for the majority of tests the null hypothesis of a unit root is not rejected. Only for the parametric bootstrap and the specification with intercepts and trends, and for the non-parametric bootstrap with intercepts the unit root hypothesis is rejected for three of the six tests. In the latter case the rejection of the null with the MW test is a borderline case with a test statistic of 310.781 and a bootstrap critical value of 309.904. Importantly, in the specification with only intercepts, the parametric and the RBB bootstrap lead to non-rejection of the unit root hypothesis for all six tests. A further important observation is that these two bootstraps indicate *incorrect* rejection of the null for three of the six tests when inference is based on the asymptotic critical values. This exemplifies again the potential pitfalls of using asymptotic critical values for the short panel at hand. Summing up, there is some evidence for unit root nonstationarity of CO_2 emissions, when using first generation panel unit root tests. Note, however, that by choosing the 'appropriate' test and by using the asymptotic critical values the rejection of the unit root null hypothesis can be 'achieved'.

We now turn to (the logarithm of real per capita) GDP. Starting with the specification including trends we see that three (parametric), two (non-parametric) and six (RBB) tests do not reject the null hypothesis of a unit root when the bootstrap critical values are used. Based on the RBB bootstrap the test decisions differ for three tests when based on the asymptotic critical values and when based on the bootstrap critical values. Thus, quite surprisingly more than for CO_2 emissions, the unit root tests lead to an unclear picture for per capita GDP. The same ambiguity prevails when including only fixed effects in the tests. Again, depending upon the choice of unit root test, bootstrap or asymptotic critical values, evidence for unit root stationarity or trend stationarity can be 'generated' by first generation panel unit root tests.

3.2 Second Generation Tests

In this subsection we now discuss the results obtained with several second generation panel unit root tests that allow for cross-sectional correlation.[15] Since there is no natural ordering in the cross-sectional dimension as compared to the time dimension, the first issue is to find tractable specifications of models for cross-sectional dependence in non-stationary panels.

[15] We do not report bootstrap inference on these second generation methods. To our knowledge an analysis of the small sample performance of these tests is still lacking. The construction of consistent bootstrap methods for cross-sectionally correlated nonstationary panels is furthermore itself an interesting question.

There are two main strands that have been followed in the literature, one is a *factor model* approach, the other is based – more classical for the panel literature – on *error components models*.

Let us turn to the idea of the factor model approach first. In this set-up the cross-sectional correlation is due to common factors that are *loaded* in all the individual country variables, e.g.

$$x_{it} = \rho_i x_{it-1} + \lambda_i' F_t + u_{it} \tag{4}$$

Here $F_t \in \mathbb{R}^k$ are the common factors and $\lambda_i \in \mathbb{R}^k$ are the so called factor loadings. In general the factors can be either stationary or integrated. After *de-factoring* the data, i.e. subtracting the factor component contained in the variables in each country, panel unit root tests (of the first generation type) can be applied to the asymptotically cross-sectionally uncorrelated de-factored data.

The most general approach in this spirit is due to Bai and Ng (2004). They provide estimation criteria for the number of factors, as well as – in the case of more than one common factor – tests for the number of common trends in the factors.[16] Thus, the factors are allowed to be stationary or integrated of order 1. After subtracting the estimated factor component, Bai and Ng (2004) propose Fisher type panel unit root tests in the spirit of Maddala and Wu (1999) and Choi (2001). The first one is asymptotically χ^2 distributed, BN_{χ^2} and the second is asymptotically standard normally distributed, BN_N. The two tests are specified against the heterogeneous alternative. See the results in Table 2. The number of common factors is estimated to be three for CO_2 and four for GDP. These estimation results are based on the information criterion BIC_3, see Bai and Ng (2004) for details. The two tests for common trends within the common factors, CT and CT_{AR}, result in three common trends, except for GDP when both fixed effects and individual trends are included (where four common trends are found).[17] Thus, essentially all common factors seem to be nonstationary. Let us next turn to the unit root tests on the de-factored data (only implemented for the fixed effects specification). Somewhat surprisingly the null hypothesis is not rejected for CO_2 emissions, but is clearly rejected for GDP by both tests. Thus, it seems that some nonstationary idiosyncratic component is present in the CO_2 emissions series.

[16]Testing for common trends can be seen as the multivariate analogue to testing for unit roots. In case of a single common factor, a unit root test for this common factor is sufficient, of course.

[17]The two tests for the number of common trends differ in the treatment of serial correlation. In CT a non-parametric correction is performed, whereas CT_{AR} is based on a vector autoregressive model fitted to the

	$NoCF$	BN_N	BN_{χ^2}	CT	CT_{AR}
Fixed Effects					
CO_2	3	-1.66	179.63	3	3
		(0.95)	(0.96)		
GDP	4	10.60	433.29	3	3
		(0.00)	(0.00)		
Fixed Effects and Trends					
CO_2	3	–	–	3	3
GDP	4	–	–	4	4

Table 2: Results of Bai and Ng (2004) PANIC analysis. $NoCF$ indicates the estimated number of common factors according to BIC_3. BN_N and BN_{χ^2} denote the unit root tests on the de-factored data. CT and CT_{AR} denote the estimated number of common trends within the common factors.
The p–values are displayed in brackets, with 0.00 indicating p–values smaller than 0.005.

Bai and Ng (2004) present the most general factor model approach to non-stationary panels currently available and the only one that allows for testing also the stochastic properties of the common factors. For completeness we also report the results obtained with two more restricted factor model approaches, due to Moon and Perron (2004) and Pesaran (2003). Moon and Perron (2004) present pooled t-type test statistics based on de-factored data (where we use the factors estimated according to Bai and Ng). We report two asymptotically standard normally distributed tests with serial correlation correction in the spirit of Phillips and Perron (1988), denoted with MP_a and MP_b. Pesaran (2003) provides an extension of the Im, Pesaran, and Shin (2003) test to allow for one factor with heterogeneous loadings. His procedure, which is a suitably cross-sectionally augmented IPS Dickey Fuller type test, works by including cross-section averages of the level and of lagged differences to the IPS-type regression. Pesaran (2003) considers two versions: the procedure just described, denoted with $C - IPS$ and a truncated, robust version $C - IPS^*$. For both of his tests the distribution is non-standard and has to be obtained by simulation methods.

The results from these factor model approaches are contained in the upper block of Table 3. The null hypothesis of a unit root is rejected in all cases (at least when testing at 6%) except for CO_2 when individual specific trends are included. Thus, all factor based unit root tests reject the unit root null hypothesis on de-factored GDP. This seems to indicate that there are global common stochastic factors (respectively trends, compare the results obtained with the

common factors.

	Fixed Effects		Fixed Effects & Trends	
	$\mathbf{CO_2}$	**GDP**	$\mathbf{CO_2}$	**GDP**
MP_a	-22.70	-17.00	-7.79	-11.58
	(0.00)	(0.00)	(0.00)	(0.00)
MP_b	-13.33	-15.70	-14.71	-27.63
	(0.00)	(0.00)	(0.00)	(0.00)
$C-IPS$	-2.09	-2.12	-1.83	-2.76
	(0.06)	(0.05)	(0.95)	(0.04)
$C-IPS^*$	-2.08	-2.11	-1.83	-2.74
	(0.06)	(0.05)	(0.95)	(0.04)
C_p	9.62	5.80	6.94	2.97
	(0.00)	(0.00)	(0.00)	(0.00)
C_Z	-8.98	-6.46	-6.79	-3.87
	(0.00)	(0.00)	(0.00)	(0.00)
C_{L^*}	-9.06	-6.15	-6.95	-3.82
	(0.00)	(0.00)	(0.00)	(0.00)
$NL-IV_1$	1.84	12.79	-0.24	-1.01
	(0.97)	(1.00)	(0.41)	(0.16)
$NL-IV_2$	8.43	13.43	0.21	-0.71
	(1.00)	(1.00)	(0.58)	(0.24)
$NL-IV_3$	3.84	11.64	0.99	1.47
	(1.00)	(1.00)	(0.84)	(0.93)

Table 3: Results of second generation panel unit root tests. The left block-column contains the results when only fixed effects are included. The right block-column contains the results when both fixed effects and individual specific linear trends are included.
In brackets the p–values are displayed, with 0.00 indicating p–values smaller than 0.005.

Bai and Ng methodology) in the GDP country data for our 107 countries. Note again that the results obtained by applying the Moon and Perron test and the Pesaran test are strictly speaking only valid if there is only one factor. For our very short panel, it may however be appropriate to compare the results obtained by several methods.

Choi (2006) presents test statistics based on an error component model. His tests are based on eliminating both the deterministic components and the cross-sectional correlations by applying cross-sectional demeaning and GLS de-trending to the data.[18] Based on these preliminary steps Choi proposes three group-mean tests based on the Fisher test principle, which differ in scaling and aggregation of the p-values of the individual tests. All three test statistics, C_p, C_Z and C_{L^*}, are asymptotically standard normally distributed. The individ-

[18]This model structure can, equivalently, be interpreted as a factor model with one factor and identical loadings for all units.

ual test underlying the implementation of this idea in the present study is the augmented Dickey-Fuller test. The results are quite clear: The null hypothesis of a unit root is rejected throughout variables and specifications.

Finally, Chang (2002) presents panel unit root tests that handle cross-sectional correlation by applying nonlinear instrumental variable estimation of the (usual) individual augmented Dickey-Fuller regressions. The instruments are given by integrable functions of the lagged levels of the variable and the test statistic is given by the standardized sum of the individual t-statistics. We present the results for three different instrument generating functions, termed $NL-IV_i$ for $i = 1, 2, 3$. The results are completely different from the other second generation panel unit root test results: The null hypothesis of a unit root is not rejected by any of the three tests for both variables and both specifications of the deterministic components. The difference in results may be explained by the Im and Pesaran (2003) comment on the Chang nonlinear IV tests. Im and Pesaran (2003) show that the asymptotic behavior established in Chang (2002) holds only for $N \ln T/\sqrt{T} \to 0$, which requires N being very small compared to T. This is of course not the case in our data set with $N = 107$ countries and $T = 13$ years. Thus, the results of the Chang NL-IV tests should be interpreted very carefully.

3.3 Conclusions from Panel Unit Root Analysis

There seems to be evidence for cross-sectional correlation for both variables. The results obtained with the method of Bai and Ng (2004) indicate the presence of three to four integrated common factors. The general conclusion from the second generation tests, except for the Chang tests, is that after subtracting the common factors, the idiosyncratic components may well be stationary. The evidence in that direction is stronger for GDP than for CO_2 emissions.

The evidence for cross-sectional correlation fundamentally weakens the basis of the results obtained by applying first generation tests. Thus, for these tests we only want to highlight again the main conclusions that can be made even without resorting to second generation methods. First, the bootstrap test distributions differ substantially from the asymptotic test distributions. This implies that test results based on bootstrap critical values can often differ from test results based on asymptotic critical values. Second, by choosing the unit root test and/or the bootstrap strategically any conclusion can be 'supported'. This large uncertainty around the results should urge researchers to be much more cautious than usual in the empirical EKC literature.

4 Panel Cointegration Tests

In this section we present panel cointegration tests for cross-sectionally uncorrelated panels. We do this to show, similarly to the panel unit root tests, that a more careful application of these methods would lead researchers to be skeptical about the validity of their results. This second level discussion is, of course overshadowed by the two first level problems.

We test for the null of no cointegration in both the linear (2) and the quadratic (1) specification of the relationship between the logarithm of per capita CO_2 emissions and the logarithm of per capita GDP. We test in quadratic version solely to show that a careful statistical analysis with the available (but inappropriate) tools of panel cointegration would already lead to ambiguous results. In particular we show that the test results depend highly upon the test applied and whether the asymptotic or some bootstrap critical values are chosen. These observations, which can be made by just using standard methods, should lead the researcher to draw only very cautious conclusions. Of course, we know from the discussion in Section 2 that cointegration in the usual sense is not defined in equation (1). This observation has been ignored in the empirical literature and several published papers, e.g., Perman and Stern (2003) discuss 'cointegration' in the quadratic specification based on unit root testing for emissions, GDP and the square of GDP.

We have in total performed ten cointegration tests, seven of them developed in Pedroni (2004) and three in Kao (1999). Similar bootstrap procedures as for the panel unit root tests are applied, see the description in Appendix B. The results obtained by applying the three tests developed by Kao are not displayed but are available from the authors upon request in a separate appendix.[19]

All tests are formulated for the null hypothesis of no cointegration, see Hlouskova and Wagner (2006b) for a discussion and a simulation based performance analysis including all

[19]Kao (1999) derives tests similar to three of the pooled tests of Pedroni for *homogenous* panels when only fixed effects are included. A panel is called homogenous, if the serial correlation pattern is identical across units. Kao's three tests, K_ρ, K_t and K_{df}, are based on the spurious least squares dummy variable (LSDV) estimator of the cointegrating regression. We have also performed these tests, since tests based on a cross-sectional homogeneity assumption might perform comparatively well even when the serial correlation patterns differ across units. This may be so, because no individual specific correlation corrections, that may be very inaccurate in short panels, have to be performed. Kao's tests are after scaling and centering appropriately asymptotically standard normally distributed and left sided. The results are qualitatively similar to the results obtained with Pedroni's tests.

the panel cointegration tests used in this paper. The tests are based on the residuals of the so called *cointegrating regression*, in our example in the linear case given by (2):[20]

$$\ln(e_{it}) = \alpha_i + \gamma_i t + \beta_1 \ln(y_{it}) + u_{it}$$

If both log emissions and log GDP are integrated, the possibility for cointegration between the two variables arises. Cointegration means that there exists a linear combination of the variables that is stationary. Thus, the null hypothesis of no cointegration in the above equation is equivalent to the hypothesis of a unit root in the residuals, $\hat{u}_{it}$ say, of the cointegrating regression. The usual specifications concerning deterministic variables have been implemented. In Table 4 we report test results when including only fixed effects and when including fixed effects and individual specific trends.

Pedroni (2004) develops four *pooled* tests and three *group-mean* tests. Three of the four pooled tests are based on a first order autoregression and correction factors in the spirit of Phillips and Ouliaris (1990). These are a variance-ratio statistic, PP_σ; a test statistic based on the estimated first-order correlation coefficient, PP_ρ; and a test based on the t-value of the correlation coefficient, PP_t. The fourth test is based on an augmented Dickey-Fuller type test statistic, PP_{df}, in which the correction for serial correlation is achieved by augmenting the test equation by lagged differenced residuals of the cointegrating regression. Thus, this test is a panel cointegration analogue of the panel unit root test of Levin, Lin, and Chu (2002). For these four tests the alternative hypothesis is stationarity with a homogeneity restriction on the first order correlation in all cross-section units.

To allow for a slightly less restrictive alternative, Pedroni (2004) develops three group-mean tests. For these tests the alternative allows for completely heterogeneous correlation patterns in the different cross-section members. Pedroni discusses the group-mean analogues of all but the variance-ratio test statistic. Similarly to the pooled tests, we denote them with PG_ρ, PG_t and PG_{df}. We report both the pooled and group-mean test results to see whether the test behavior differs systematically between these two types of tests.

After centering and scaling the test statistics by suitable correction factors, to correct for serial correlation of the residuals and for potential endogeneity of the regressors in the cointegrating regression, all test statistics are asymptotically standard normally distributed.

[20] For such a short panel as given here, systems based methods like the one developed in Groen and Kleibergen (2003) are not applicable.

Figures similar to Figure 1 are available from the authors upon request. Again substantial differences between the asymptotic critical values and the bootstrap critical values emerge.

The first block in Table 4 corresponds to the parametric bootstrap, the second to the non-parametric bootstrap and the third to the RBB bootstrap. Within each block, the first block-row corresponds to the linear specification and the second to the quadratic specification. Both, the linear and the quadratic specification have been tested with fixed effects and with fixed effects and individual specific linear trends. Just to be sure, note again, that testing for cointegration in the quadratic formulation lacks theoretical econometric foundations.

Let us start with the linear specification, which is 'only' subject to the first level problem of cross-sectional correlation. There is some variability of results across bootstrap methods and again in a variety of cases bootstrap inference leads to different conclusions than resorting to the asymptotic critical values. This happens in particular for the RBB bootstrap. For the quadratic specification, i.e. the Kuznets curve in its usual formulation, roughly the same *observations* as for the linear specification can be made, ignoring again the problem that a correct econometric foundation is lacking due to the nonlinear transformation. Again the RBB bootstrap leads to the fewest rejections of the null hypothesis. The null hypothesis of no cointegration is more often rejected for the linear formulation than for the quadratic specification. Note that no systematic differences between the pooled and the group-mean tests occur.

The above results provide some weak evidence for the presence of a cointegrating relationship between GDP and emissions. However, as for the panel unit root tests, by choosing the test and the bootstrap strategically, any 'conclusion' can be supported. This 'volatility' of the results should lead researchers to be more cautious than what is usually observed.

	PP_σ	PP_ρ	PP_t	PP_{df}	PG_ρ	PG_t	PG_{df}
PARAMETRIC BOOTSTRAP							
Linear Specification							
FE	**2.887*** (3.670)	**-3.121** (-3.129)	**-7.061** (-6.501)	**-4.899** (-0.041)	**-1.595** (-0.041)	**-10.498** (-6.569)	**-10.956** (-6.949)
FE & Tr.	1.155 (1.484)	**-3.045** (-0.872)	**-14.877** (-9.721)	**-12.827** (-7.840)	**1.477** (3.002)	**-13.058** (-8.046)	**-14.352** (-9.149)
Quadratic Specification							
FE	0.434 (2.761)	-0.382 (-1.197)	**-6.528*** (-7.827)	**-4.769*** (-6.173)	**2.335** (2.401)	**-8.766** (-8.035)	**-9.718** (-8.927)
FE & Tr.	-0.954 (-0.457)	0.639 (2.174)	**-14.454** (-9.529)	**-11.985** (-7.456)	**4.592** (5.611)	**-13.325** (-9.386)	**-14.596** (-10.554)
NON-PARAMETRIC BOOTSTRAP							
	PP_σ	PP_ρ	PP_t	PP_{df}	PG_ρ	PG_t	PG_{df}
Linear Specification							
FE	**2.887** (0.858)	**-3.121** (-1.880)	**-7.061** (-5.261)	**-4.899** (-3.593)	**-1.595** (0.364)	**-10.498** (-6.124)	**-10.956** (-6.424)
FE & Tr.	1.155 (1.498)	**-3.045** (-0.814)	**-14.877** (-9.582)	**-12.827** (-7.772)	**1.477** (3.027)	**-13.058** (-7.873)	**-14.352** (-9.018)
Quadratic Specification							
FE	0.434 (1.023)	-0.382 (-0.792)	**-6.528*** (-7.193)	**-4.769*** (-5.574)	**2.336** (2.647)	**-8.766** (-7.552)	**-9.718** (-8.405)
FE & Tr.	-0.953 (-0.413)	**0.639** (2.160)	**-14.454** (-9.429)	**-11.985** (-7.432)	**4.592** (5.586)	**-13.325** (-9.331)	**-14.596** (-10.562)
RESIDUAL BASED BLOCK BOOTSTRAP							
	PP_σ	PP_ρ	PP_t	PP_{df}	PG_ρ	PG_t	PG_{df}
Linear Specification							
FE	**2.887*** (3.089)	**-3.121** (-2.947)	**-7.061** (-6.933)	**-4.899*** (-5.269)	**-1.595** (-0.915)	**-10.498** (-9.134)	**-10.956** (-9.777)
FE & Tr.	1.155 (2.947)	**-3.045*** (-3.296)	**-14.877** (-15.231)	**-12.827** (-12.725)	1.477 (0.936)	**-13.058*** (-14.977)	**-14.352*** (-16.341)
Quadratic Specification							
FE	0.434 (0.975)	-0.382 (-1.274)	**-6.528*** (-8.733)	**-4.769*** (-7.016)	2.336 (1.085)	**-8.766*** (-12.976)	**-9.718*** (-14.153)
FE & Tr.	-0.936 (0.503)	0.639 (0.0490)	**-14.454** (-15.303)	**-11.985*** (-12.150)	4.592 (3.604)	**-13.325*** (-17.100)	**-14.596*** (-19.070)

Table 4: Results of Pedroni's panel cointegration tests including fixed effects only (FE), respectively fixed effects and linear trends (FE & Tr.). The upper block-row displays the results for the linear specification and the lower block-row displays the results for the quadratic specification. The first part of the table corresponds to the parametric bootstrap, the second to the non-parametric bootstrap and the third to the residual based block bootstrap.
The asymptotic 5 % critical value is given by 1.645 for the first test and by -1.645 for the other 6 tests.
Bold indicates rejection of the null hypothesis based on the bootstrap critical values and **bold*** indicates rejection based upon the asymptotic critical values but no rejection according to the bootstrap critical values.
The autoregressive lag lengths in both the autoregression based tests, the parametric bootstrap and the non-parametric bootstrap are equal to 1. The window-length of the Bartlett kernels used in the non-parametric tests is also equal to 1. The block-length in the RBB bootstrap is equal to 2.

5 Estimation of the Carbon Kuznets Curve with Panel Cointegration Methods and Using De-factored Observations

We finally turn to estimating the CKC relationship. In the first subsection we estimate the CKC with panel cointegration methods that correspond to the first generation panel unit root and cointegration tests. These methods are of course subject to the two first level critiques. As for the panel unit root and cointegration tests, we include results based on this type of methods to show that by careful application the conclusions one could draw, even when staying in this framework, are very weak. In the second subsection we estimate the CKC relationship on de-factored data. These are, up to potentially bad small sample performance of the Bai and Ng (2004) procedure, stationary. Thus, for these data standard panel regression techniques are applicable. Note also that the de-factored data are (asymptotically) cross-sectionally uncorrelated.

5.1 Panel Cointegration Estimation

Two types of estimators for the cointegrating relationship in panels are applied: *fully modified* ordinary least squares (FM-OLS) and *dynamic* ordinary least squares (D-OLS). Both estimation methods are panel extensions of well known time series concepts. FM-OLS was introduced by Phillips and Hansen (1990) and D-OLS is due to Saikkonen (1991). Both methods allow for serial correlation in the residuals and for endogeneity of the regressors in the cointegrating regression. The panel extensions of FM-OLS are discussed in detail in Phillips and Moon (1999), nesting the discussions in Pedroni (2000) and Kao and Chiang (2000). As in the time series case, the idea of FM-OLS is to obtain in the first step OLS estimates of long-run variance matrices. In the second step another regression is run on *corrected* variables, with the correction factors being functions of the estimated long-run variance matrices. The idea of D-OLS is to correct for serial correlation and endogeneity by augmenting the cointegrating regression by leads and lags of first differences of the regressors. The panel extensions of D-OLS are discussed in Kao and Chiang (2000) and Mark and Sul (2003). Both methods, FM-OLS and D-OLS, yield asymptotically normally distributed (for first T followed by N to infinity) estimated cointegrating vectors, which implies that χ^2 inference via e.g. Wald tests can be conducted. Note for completeness that various versions of both FM-OLS and D-OLS in weighted or unweighted fashions have been implemented, see Hlouskova and Wagner (2006b) for a description. These differ i.a. in how the correction factors are computed.

	Fixed Effects			
	FM-OLS	D-OLS	wD-OLS	LSDV
$\ln y_{it}$	0.461	1.401	0.478	0.508
	(23.358)	(4.431)	(14.119)	(3.948)
$(\ln y_{it})^2$	0.046	-0.030	0.216	0.014
	(2.3221)	(-1.338)	(6.387)	(1.5797)
	Fixed Effects and Trends			
	FM-OLS	D-OLS	wD-OLS	LSDV
$\ln y_{it}$	0.341	1.860	0.663	0.239
	(17.282)	(8.969)	(19.584)	(1.252)
$(\ln y_{it})^2$	0.208	-0.092	0.205	0.012
	(10.548)	(-5.805)	(6.069)	(0.855)

Table 5: Estimation results for equation (1) including fixed effects only in the upper block and fixed effects and linear trends in the lower panel. Fixed effects, respectively fixed effects and trend slopes are not reported. In brackets the t-statistics are displayed.
Note that the results (coefficients and t-values) in this table do not have a theoretical underpinning due to the use of nonlinear transformations of integrated processes.

Let us start with a discussion of the results obtained when estimating the linear formulation (2). Note again that the linear specification is 'only' subject to the problem of cross-sectional correlation, i.e. only to one of the first level problems. In the specification including only fixed effects, the coefficient of log per capita GDP is between 0.6 and 0.8, depending upon estimation method. For the specification including unit specific trends, the estimated coefficient on log per capita GDP varies between 0.4 and 0.8, depending upon estimation method. The null hypothesis of a unit GDP elasticity of emissions, i.e the null hypothesis $H_0 : \beta_1 = 1$ in equation (2), is rejected for all estimation methods and specifications.

We now turn to the estimation results obtained for the quadratic formulation (1), which is subject to both first level problems. Table 5 reports one FM-OLS estimation result and two different versions of D-OLS estimation results, abbreviated by D-OLS and wD-OLS, due to Mark and Sul (2003) and Kao and Chiang (2000). We report two different D-OLS results to show that various D-OLS implementations deliver substantially differing results. For the FM-OLS estimates less variability across versions occurs than for the D-OLS estimates. Thus, only the results of one FM-OLS variant are reported. Important in this respect is the observation that such a large variability of estimated coefficients across methods might already by itself indicate underlying problems. The results obtained by applying the D-OLS estimator of Kao and Chiang (2000) are very different from the rather similar FM-OLS and wD-OLS estimation results. Only the D-OLS estimates have a *negative* coefficient for squared log GDP. Thus, only the results derived with this estimator imply an inverse U–shape. The 'turning point' of these inverse U–shapes, however, is highly implausible. It is at about 17.3 million dollars for the fixed effects case and at about 220 dollars for the fixed effects and trends case. Both numbers are neither sensible nor useful and should lead to reconsider the usefulness of the estimation methods for the problem at hand (or the usefulness of the specification).

The final column in Table 5 reports the estimation results based on the LSDV estimator, to see which kind of results are obtained when ignoring the nonstationarity issue altogether. When only fixed effects are included, the difference to the FM-OLS and wD-OLS estimators are not too large. However, when fixed effects and trends are included, the differences to the cointegration results become substantial. Furthermore, no coefficient appears to be significant in that case. By choosing other estimators for stationary panels, all kinds of results can be generated. Thus, also when ignoring issues of nonstationarity a researcher can or cannot come to the conclusion of the prevalence of a relationship between emission and GDP, depending upon the specification of the deterministic component and the estimation method.

5.2 Estimation with De-Factored Observations

We finally report the estimation results based on the de-factored observations, using the approach developed by Bai and Ng (2004) for de-factoring the data. Remember from Section 3 that three respectively four common factors have been found, all of which seem to be nonstationary, according to the Bai and Ng tests. An application of the unit root tests of Bai and Ng (2004) to the de-factored data indicates that the idiosyncratic components are stationary. This implies that for the de-factored data standard regression theory developed for stationary variables applies. The results are displayed in Table 6. We present two estimation results. The first applies if de-factorization is performed in the model with only fixed effects ($DF-2$) and the second when de-factorization is performed in the model with fixed effects and trends ($DF-3$). The preferred specifications of the estimated CKCs contain in both cases fixed country and time effects.[21] GLS estimation with cross section weights is performed to allow for different error variances across countries.

Since the data are de-factored here, the size of the coefficients cannot be directly compared with the results of Table 5, ignoring for the moment that the results presented in Table 5 are

[21] In the first case, when including trends in the regression, significant coefficients emerge for some countries. However, the specification with time effects is preferred.

	$DF-2$	$DF-3$
$\ln y_{it}$	0.389	0.472
	(6.223)	(6.961)
$(\ln y_{it})^2$	1.130	3.290
	(1.830)	(3.566)

Table 6: Estimation results for equation (1) on de-factored data. Estimation is performed by GLS. In brackets robust t-statistics are displayed.

subject to the problems discussed throughout the paper. Both coefficients are positive and significant, the coefficient on squared log per capita GDP in $DF-2$ only at 7%. Thus, there is no evidence for an inverse U–shaped relationship as postulated by the CKC hypothesis. Of course, these results are subject to the properties of de-factorization for short samples, which are yet not well understood. Apart from this problem, however, these estimates are the only ones presented in this paper that are based on an asymptotically well founded estimation theory, given that the data are indeed unit root nonstationary. Therefore, with all reservation necessary, we tentatively conclude that within our panel data set no evidence for an inverse U–shape relation between log per capita GDP and log per capita CO_2 emissions is present (after de-factoring the data).

6 Summary and Conclusions

In this paper we discuss three important econometric problems associated with the Environmental Kuznets Curve, that arise when the data are of the unit root nonstationary type. We exemplify the discussion for the Carbon Kuznets Curve, relating per capita GDP to per capita emissions of CO_2, on a panel comprising 107 countries over the years 1986–1998.

The three problems are grouped in two first level problems and one second level problem. The two first level problems are the use of nonlinear transformations of integrated processes as regressors and cross-sectional dependence in nonstationary panels. The second level problem is the poor performance of (panel) unit root and cointegration techniques for short time series or panels.

Let us start with the first level problems. The discussion in Section 2 shows that nonlinear transformations – like the square – of an integrated process are in general not integrated. This implies that the usual unit root and cointegration techniques cannot be applied for the EKC and CKC, if log per capita GDP is indeed integrated. This point has been completely overlooked in the empirical EKC and CKC literature up to now, even in that part of the literature that acknowledges the potential presence of integrated processes. We do not solve the problem in this study, since up to now no estimation techniques for panels containing nonlinear transformations of integrated processes are available. Currently only results for the time series case, developed by Peter Phillips and co-authors, are available. Ongoing research is investigating the applicability of (panel extensions of) these methods to EKC/CKC estimation.

To address the second of the first level problems, cross-sectional dependence in nonstationary panels, the literature offers several approaches in the meantime. Prior to this study, only so called first generation panel unit root and cointegration techniques have been applied, which all rely upon cross-sectional independence. In the CKC case this amounts to independence of both GDP and CO_2 emissions across countries. We present in this paper the first application of second generation methods that allow for cross-sectional correlation in the EKC/CKC context. The results obtained with the method of Bai and Ng (2004) indicate that non-stationary common factors may well be present in both GDP and emissions. The results also indicate that the idiosyncratic components (i.e. the de-factored data) are stationary. In this respect the evidence is stronger for GDP than for emissions. Based on these findings we estimate the CKC on de-factored data, which are cross-sectionally uncorrelated and, see above, also stationary. Thus, standard panel regression techniques are applicable to the de-factored data and also the nonlinearly transformed regressor does not pose additional problems in the stationary context. We find no evidence for an inverse U–shape relationship. These results are, of course, subject to potentially bad small sample performance of the Bai and Ng de-factoring procedure, potential failure of the homogeneity assumption across countries and potential structural instabilities over time. The first issue is not yet understood in practice and the second and third issue have not been discussed in detail in this paper, since the focus in this paper is solely on the implications of unit root nonstationarity on the estimation of Environmental Kuznets Curves.

The second level problem is the, in our opinion, relatively uncritical use of unit root and cointegration methods in the EKC/CKC literature. It is known that unit root and cointegration techniques perform poor for short time series. This poor performance translates into poor performance for short panels, see Hlouskova and Wagner (2004a,b) for simulation evidence. Staying within the first generation framework (and thus ignoring the first level problems!), we show that careful application of the methods indicates that the results should be interpreted with caution. By implementing three different bootstrap algorithms we show that (three different estimates of) the finite sample distributions differ substantially from the asymptotic distributions. This implies that inference based on the asymptotic critical values can be highly misleading. Thus, we conclude that by 'strategic' choice of the unit root and cointegration tests any conclusion can be 'supported'. This holds, to a lesser extent, even when resorting to bootstrapping, where the RBB bootstrap results differ in several cases from the other two. This finding, however, may be due to the short time dimension that poses a challenge to block re-sampling based bootstrap schemes. The results for the two other bootstrap algorithms are rather similar.

Ignoring the first level problems also for estimation, we estimate the CKC with panel cointegration estimators. This exercise leads to highly variable results across different variants of estimators, with less variability across the FM-OLS variants than across the D-OLS variants. From this variability we conclude that also estimation results obtained *within* the first generation framework should have been interpreted with much more caution than has been done in the literature.

Summing up we conclude – a bit polemically – that a large part of the empirical EKC and CKC literature up to now has been plagued by using inappropriate methods in a sloppy manner. Hence, the title of the paper. However, recent progress made in the theoretical literature will soon equip the empirical researcher with the necessary tools to clear the sky.

References

Bai, J. and S. Ng (2004). A PANIC Attack on Unit Roots and Cointegration. *Econometrica* **72**, 1127–1178.

Bertinelli, L. and E. Strobl (2005). The Environmental Kuznets Curve Semi-Parametrically Revisited. *Economics Letters* **88**, 350–357.

Bradford, D., R. Fender, S.H. Shore, and M. Wagner (2005). The Environmental Kuznets Curve: Exploring a Fresh Specification. *Contributions to Economic Analysis and Policy* **4**, No. 1, Article 5. Berkeley Electronic Press.

Breitung, J. (2000). The Local Power of some Unit Root Tests for Panel Data, 161–177. In Baltagi, B.H. (Ed.) *Nonstationary Panels, Panel Cointegration, and Dynamic Panels*, Elsevier, Amsterdam.

Chang, Y., J.Y. Park, and P.C.B. Phillips (2001). Nonlinear Econometric Models with Cointegrated and Deterministically Trending Regressors. *Econometrics Journal* **4**, 1–36.

Chang, Y. (2002). Nonlinear IV Unit Root Tests in Panels with Cross-Sectional Dependency. *Journal of Econometrics* **110**, 261–292.

Choi, I. (2001). Unit Root Tests for Panel Data. *Journal of International Money and Finance* **20**, 249–272.

Choi, I. (2006). Combination Unit Root Tests for Cross-Sectionally Correlated Panels, 311–333. In Corbae, D., S. Durlauf and B. Hansen (Eds.) *Econometric Theory and Practice: Frontiers of Analysis and Applied Research*, Cambridge University Press, Cambridge.

Dijkgraaf, E. and H.R.J. Vollebergh (2005). A Note on Testing for Environmental Kuznets Curves with Panel Data. *Environmental and Resource Economics* **32**, 229–239.

Fisher, R.A. (1932). *Statistical Methods for Research Workers.* Oliver and Boyd, Edinburgh.

Groen, J.J.J. and F. Kleibergen (2003). Likelihood-Based Cointegration Analysis in Panels of Vector Error-Correction Models. *Journal of Business and Economic Statistics* **21**, 295–318.

Grossmann, G.M. and A.B. Krueger (1991). Environmental Impacts of a North American Free Trade Agreement. NBER Working paper No. 3914.

Grossmann, G.M. and A.B. Krueger (1993). Environmental Impacts of a North American Free Trade Agreement, 13–56. In Garber, P. (Ed.) *The Mexico-US Free Trade Agreement*, MIT Press, Cambridge.

Grossmann, G.M. and A.B. Krueger (1995). Economic Growth and the Enivronment. *Quartely Journal of Economics* **110**, 353–377.

Harris, R.D.F. and E. Tzavalis (1999). Inference for Unit Roots in Dynamic Panels Where the Time Dimension is Fixed. *Journal of Econometrics* **90**, 1–44.

Heil, M.T. and T.M. Selden (1999). Panel Stationarity with Structural Breaks: Carbon Emissions and GDP. *Applied Economics Letters* **6**, 223–225.

Hlouskova, J. and M. Wagner (2006a). The Performance of Panel Unit Root and Stationarity Tests: Results from a Large Scale Simulation Study. *Econometric Reviews* **25**, 85–116.

Hlouskova, J. and M. Wagner (2006b). The Performance of Panel Cointegration Methods: Results from a Large Scale Simulation Study. Mimeo.

Holtz-Eakin, D. and T.M. Selden (1995). Stoking the Fires? CO_2 Emissions and Economic Growth. *Journal of Public Economics* **57**, 85–101.

Im, K.S. and M.H. Pesaran (2003). On the Panel Unit Root Tests Using Nonlinear Instrumental Variables. Mimeo, University of Cambridge.

Im, K.S., M.H. Pesaran, and Y. Shin (1997). Testing for Unit Roots in Heterogeneous Panels. Mimeo.

Im, K.S., M.H. Pesaran, and Y. Shin (2003). Testing for Unit Roots in Heterogeneous Panels. *Journal of Econometrics* **115**, 53–74.

Kao, C. (1999). Spurious Regression and Residual Based Tests for Cointegration in Panel Data. *Journal of Econometrics* **90**, 1–44.

Kao, C. and M.-H. Chiang (2000). On the Estimation and Inference of a Cointegrated Regression in Panel Data. In Baltagi, B.H. (Ed.) *Nonstationary Panels, Panel Cointegration, and Dynamic Panels*, Elsevier, Amsterdam.

Lanne, M. and M. Lisky (2004). Trends and Breaks in Per-Capita Carbon Dioxide Emissions, 1870–2028. *Energy Journal* **25**, 41–65.

Levin, A., C.F. Lin, and C-S.J. Chu (2002). Unit Root Tests in Panel Data: Asymptotic and Finite Sample Properties. *Journal of Econometrics* **108**, 1–22.

Maddala, G.S. and S. Wu (1999). A Comparative Study of Unit Root Tests with Panel Data and a Simple New Test. *Oxford Bulletin of Economics and Statistics* **61**, 631–652.

Manne, A.S, R.O. Mendelsohn, and R.G. Richels (1995). MERGE: A Model for Evaluating Regional and Global Effects of GHG Reduction Policies. *Energy Policy* **23**, 17–34.

Mark, N.C. and D. Sul (2003). Cointegration Vector Estimation by Panel Dynamic OLS and Long-Run Money Demand. *Oxford Bulletin of Economics and Statistics* **65**, 655–680.

Martinez-Zarzoso, I. and A. Bengochea-Morancho (2004). Pooled Mean Group Estimation of an Environmental Kuznets Curve for CO_2. *Economics Letters* **82**, 121–126.

McKibbin, W., M. Ross, R. Shackleton, and P. Wilcoxen (1999). Emissions Trading, Capital Flows and the Kyoto Protocol. *Energy Journal*, special volume edited by Weyant, J.P.: *The Costs of the Kyoto Protocol: A Multi-Model Evaluation*, 287–334.

Millimet, D.L., J.A. List, and T. Stengos (2003). The Environmental Kuznets Curve: Real Progress or Misspecified Models? *Review of Economics and Statistics* **85**, 1038–1047.

Moon, H.R. and B. Perron (2004). Testing for a Unit Root in Panels with Dynamic Factors. *Journal of Econometrics* **122**, 81–126.

Müller-Fürstenberger G. and M. Wagner (2006). Exploring the Environmental Kuznets Hypothesis: Theoretical and Econometric Problems. Forthcoming in *Ecological Economics*.

Nordhaus, W. (1992). An Optimal Transition Path for Controlling Greenhouse Gases. *Science* **228**, 1315–1319.

Paparoditis, E. and D. Politis (2003). Residual-based Block Bootstrap for Unit Root Testing. *Econometrica* **71**, 813–855.

Paparoditis, E. and D. Politis (2005). Bootstrapping Unit Root Tests for Autoregressive Time Series. *Journal of the American Statistical Association* **100**, 545–553.

Park, J.Y. and P.C.B. Phillips (1999). Asymptotics for Nonlinear Transformations of Integrated Time Series. *Econometric Theory* **15**, 269–298.

Park, J.Y. and P.C.B. Phillips (2001). Nonlinear Regressions with Integrated Time Series. *Econometrica* **69**, 117–161.

Pedroni, P. (2000). Fully Modified OLS for Heterogeneous Cointegrated Panels. In Baltagi, B.H. (Ed.) *Nonstationary Panels, Panel Cointegration, and Dynamic Panels*, Elsevier, Amsterdam.

Pedroni, P. (2004). Panel Cointegration. Asymptotic and Finite Sample Properties of Pooled Time Series Tests with an Application to the PPP Hypothesis. *Econometric Theory* **20**, 597–625.

Peixoto, J.P. and A.H. Oort (1992). Physics of the Climate. *American Institute of Physics*, New York, 520pp.

Perman, R. and D.I. Stern (2003). Evidence from Panel Unit Root and Cointegration Tests that the Environmental Kuznents Curve does not exist. *Australian Journal of Agricultural and Resource Economics* **47**, 325–347.

Pesaran, M.H. (2003). A Simple Panel Unit Root Test in the Presence of Cross Section Dependence. Mimeo, University of Cambridge.

Phillips, P.C.B. (1986). Understanding Spurious Regressions in Econometrics. *Journal of Econometrics* **33**, 311–340.

Phillips, P.C.B. and B.E. Hansen (1990). Statistical Inference in Instrumental Variables Regression with I(1) Processes. *Review of Economic Studies* **57**, 99–125.

Phillips, P.C.B. and H.R. Moon (1999). Linear Regression Limit Theory for Nonstationary Panel Data. *Econometrica* **67**, 1057–1111.

Phillips, P.C.B. and S. Ouliaris (1990). Asymptotic Properties of Residual Based Tests for Cointegration. *Econometrica* **58**, 165–193.

Phillips, P.C.B. and P. Perron (1988). Testing for a Unit Root in Time Series Regression. *Biometrika* **75**, 335–346.

Ramanathan, V., R.J. Cicerone, H.B. Singh, and J.T. Kiehl (1985). Trace Gas Trends and their Potential Role in Climate Change. *Journal of Geophysical Research* **90**, 5547–5566.

Saikkonen, P. (1991). Asymptotically Efficient Estimation of Cointegrating Regressions. *Econometric Theory* **7**, 1–21.

Schmalensee, R., T.M. Stoker, and R.A. Judson (1998). World Carbon Dioxide Emissions: 1950–2050. *Review of Economics and Statistics* **80**, 15–27.

Stern, D.I. (2004). The Rise and Fall of the Environmental Kuznets Curve. *World Development* **32**, 1419–1439.

Yandle, B., M. Bjattarai, and M. Vijayaraghavan (2004). Environmental Kuznets Curves: A Review of Findings, Methods, and Policy Implications. Research Study 02.1 update, PERC.

Appendix: Data and Sources

Our analysis is based on balanced panel data for 107 countries for the period 1986–1998 listed in Table 7. The former Soviet Union and some eastern European countries are omitted from the sample because of missing data. Other countries like Kuwait are omitted because of large jumps in the emissions data.

Per capita CO_2 emissions are taken from the Carbon Dioxide Information Analysis Center (CDIAC) data set (see `http://cidia.eds.ornl.gov/trends/emis/emcont.html`). They are measured in metric tons of CO_2. We transform them to kilograms to achieve variables of comparable magnitude as the per capita GDP series. Per capita GDP is measured in constant 1995 US$ and taken from the World Bank Development Indicators 2003.

Albania	Ecuador	Liberia	Seychelles
Algeria	Egypt	Luxembourg	Singapore
Antigua Barbuda	El Salvador	Macao	Solomon Islands
Argentina	Fiji	Malaysia	South Africa
Australia	Finland	Malta	Spain
Austria	France	Mauritania	Sri Lanka
Bahamas	French Guiana	Mauritius	St. Lucia
Bahrain	Gabon	Mexico	St. Vincent and Grenadines
Barbados	Germany	Mongolia	Suriname
Belgium	Greece	Morocco	Swaziland
Belize	Grenada	Netherlands	Sweden
Bolivia	Guatemala	New Caledonia	Switzerland
Botswana	Guyana	New Zealand	Syrian Arab. Rep.
Brazil	Honduras	Nicaragua	Thailand
Brunei	Hong Kong	Nigeria	Tonga
Bulgaria	Hungary	Norway	Trinidad and Tobago
Cameroon	Iceland	Oman	Tunisia
Canada	India	Pakistan	Turkey
Chile	Indonesia	Panama	United Arab. Emirates
China	Iran	Papua New Guinea	United Kingdom
Colombia	Ireland	Paraguay	United States
Costa Rica	Israel	Peru	Uruguay
Cyprus	Italy	Philippines	Venezuela
Denmark	Jamaica	Portugal	Vietnam
Djibouti	Japan	Puerto Rico	Zambia
Dominica	Jordan	Romania	Zimbabwe
Dominican Rep.	Korea Rep.	Saudi Arabia	

Table 7: List of countries included in the computations.

Appendix B: Bootstrap Algorithms

Bootstrapping the first generation panel unit root and panel cointegration tests used in this paper requires to take two issues into consideration. First, unit root nonstationarity of certain quantities (all tests applied have the null of a unit root in the panel, and correspondingly of no cointegration). Second, the serial correlation allowed for in the innovation processes u_{it}.

Both issues can be handled by resorting to appropriate bootstrap procedures. Bootstrap procedures for unit root nonstationary processes are in the meantime relatively well understood, see e.g. Paparoditis and Politis (2003). In our application we have to take into account in addition the small time dimension of our panels. For this reason, one part of our bootstrap procedures fits an autoregression to the residuals of the unit root test equation respectively of the cointegrating regression. Bootstrapping is then based on the residuals from these autoregressive approximations, which resemble white noise. For our case with $T = 13$ this might be preferable to some block-bootstrap procedure. For comparison, however, we have also implemented the so called *residual based block* bootstrap (RBB) procedure of Paparoditis and Politis (2003), which has certain asymptotical (for $T \to \infty$) advantages in terms of power compared to the other procedures implemented, compare Paparoditis and Politis (2005).

In the panel case we have to consider bootstrapping in such a way that cross-sectional correlation is preserved. A simple way of achieving this is to re-sample the residuals with the same re-sampling schemes for all units. In this respect the simulation results of Hlouskova and Wagner (2006a, 2006b) indicate that the tests are robust to a certain amount of short-run dependence. Note, however, that none of the first generation tests has been designed for cross-sectionally correlated panels.

Note that the panel unit root tests and panel cointegration tests are implemented for two different specifications concerning the deterministic components. One, where only fixed effects are contained in the test equation respectively the cointegrating regression and the other where both fixed effects and individual trends are contained. We only discuss the second case in this appendix, the other case follows trivially.

Let us now discuss the bootstrapping algorithms implemented for the panel unit root tests and let us start with the autoregression based algorithms. Denote with $y_{it} \in \mathbb{R}$ the panel data observed for $i = 1, \dots, N$ and $t = 1, \dots, T$. Then for each unit the following equation is estimated by OLS:

$$\Delta y_{it} = \gamma_{i0} + \sum_{j=1}^{p_i} \gamma_{ij} \Delta y_{it-j} + u_{it} \tag{5}$$

with Δ denoting the first difference operator. The lag lengths p_i are allowed to vary across the individual units in order to whiten the residuals u_{it}. Denote with $\hat{u}_{it}$ the residuals of equation (5). Then the following two bootstrap procedures are based on the autoregression residuals.

(i) Parametric: The bootstrap residuals are given by $u^*_{it} = \hat{\sigma}_i \varepsilon_{it}$, where $\hat{\sigma}^2_i$ denotes the estimated variance of $\hat{u}_{it}$ and $\varepsilon_{it} \sim N(0,1)$.

(ii) Non-parametric:[22] Denote with $\hat{u}_t = \begin{bmatrix} \hat{u}_{1t}, & \ldots, & \hat{u}_{Nt} \end{bmatrix}'$ and generate the bootstrap residuals u^*_t by re-sampling $\hat{u}_t, t = p+2, \ldots, T$ with replacement. By re-sampling the whole vector, contemporaneous correlation across units (stemming from the residuals) is preserved in the bootstrap series.

Given u^*_{it} the bootstrap data themselves are generated from

$$y^*_{it} = \begin{cases} y_{it} & t = 1, \ldots, p_i + 1 \\ \hat{\gamma}_{i0} + y^*_{it-1} + \sum_{j=1}^{p_i} \hat{\gamma}_{ij} \Delta y^*_{it-j} + u^*_{it} & t = p_i + 2, \ldots, T \end{cases} \tag{6}$$

As indicated above Paparoditis and Politis (2003) propose a different bootstrap algorithm, the RBB bootstrap, based on *unrestricted* residuals. By unrestricted residuals we mean residuals which are not generated from an equation like (5) where a unit root is imposed, due to estimation in first differences, but from an unrestricted first order autoregression. Higher order serial correlation is not dealt with by fitting an autoregression, but by bootstrapping blocks, with the block-length increasing with sample size at a sufficient rate.[23] The implementation of the RBB bootstrap is as follows:

(i) Estimate the equation $y_{it} = \gamma_{i0} + \rho_i y_{it-1} + u_{it}$ by OLS (for each unit).

(ii) Calculate the centered residuals

$$\tilde{u}_{it} = (y_{it} - \hat{\rho}_i y_{it-1}) - \frac{1}{T-1} \sum_{\tau=2}^{T} (y_{i\tau} - \hat{\rho}_i y_{i\tau-1}).$$

[22] For notational simplicity we assume $p_i = p$ for all units here in the discussion.

[23] For an autoregression based implementation of this idea of using unrestricted residuals see Paparoditis and Politis (2005).

(iii) Choose the block-length b and draw $j_0, \ldots, j_{k-1}$ from the discrete uniform distribution over the set $\{1, \ldots, T-b\}$ with $k = \lfloor \frac{T-1}{b} \rfloor$. Here $\lfloor x \rfloor$ denotes the integer part of x. By taking the same realizations j_m for all cross-sections, the cross-sectional correlation is preserved in the bootstrap data.

(iv) Denoting with $m = \lfloor \frac{t-2}{b} \rfloor$ and with $s = t - mb - 1$, the bootstrap data are given by:

$$y^*_{it} = \begin{cases} y_{i1} & t = 1 \\ \hat{\gamma}_{i0} + y^*_{it-1} + \tilde{u}_{ij_m+s} & t = 2, \ldots, kb+1 \end{cases} \tag{7}$$

Note again for completeness that for the tests that only allow for an intercept in the test equation $\hat{\gamma}_{i0}$ above is replaced by zero.

For the panel cointegration tests used in this study we also apply three bootstrap algorithms. These are essentially multivariate extensions of the above. The starting point for the autoregression based bootstrap procedures is now given by

$$y_{it} = \alpha_i + \delta_i t + X'_{it}\beta_i + u_{it} \tag{8}$$

$$X_{it} = A_i + X_{it-1} + \varepsilon_{it} \tag{9}$$

for $i = 1, \ldots, N$, $t = 1, \ldots, T$. Now $\alpha_i, \delta_i \in \mathbb{R}$, $X_{it} = [x_{it1}, \ldots, x_{itk}]'$ and $A_i, \beta_i \in \mathbb{R}^k$. Note for completeness that for the test proposed by Kao (1999) $\beta_i = \beta$ holds for all units. Under the null hypothesis of no cointegration between y_{it} and X_{it} it follows that u_{it} is integrated and that ε_{it} is stationary.

We estimate[24] the above equations (8) and (9) to obtain the estimated residuals $\hat{v}_{it} = [\hat{u}_{it}, \hat{\varepsilon}'_{it}]'$ from

$$\begin{aligned} \hat{u}_{it} &= y_{it} - \hat{\alpha}_i - \hat{\delta}_i t - X'_{it}\hat{\beta}_i \\ \hat{\varepsilon}_{it} &= \Delta X_{it} - \hat{A}_i \end{aligned}$$

Under the null hypothesis $v_{it} \in \mathbb{R}^{k+1}$ is a process whose first coordinate is integrated and whose other coordinates are stationary. These known restrictions can be incorporated into the autoregressive modelling to obtain white noise residuals by fitting a vector error correction model which incorporates the exact knowledge about the cointegrating space. This is achieved by estimating:

$$\hat{v}_{it} = B_i \hat{\varepsilon}_{it-1} + \sum_{j=1}^{p_i} \Gamma_j \Delta \hat{v}_{it-j} + \mu_{it} \tag{10}$$

[24] Estimation proceeds by unit specific OLS estimation, except for the method of Kao (1999), which rests upon the LSDV estimator to obtain an estimate $\hat{\beta}$ identical across units.

with $B_i \in \mathbb{R}^{k+1\times k}$. The residuals from equation (10), $\hat{\mu}_{it}$ say, resemble white noise due to appropriate choice of the lag lengths p_i.

As in the univariate case for the panel unit root tests, two bootstrap versions are implemented based on $\hat{\mu}_{it}$.

(i) Parametric: Estimate the variance-covariance matrix of $\hat{\mu}_{it}$, $\hat{\Sigma}_i$ say. Denote its lower triangular Cholesky factor by $\hat{L}_i$ and generate the bootstrap residuals $\mu^*_{it} = \hat{L}_i \eta_{it}$ with $\eta_{it} \sim N(0, I_{k+1})$.

(ii) Non-parametric: μ^*_{it} is given by re-sampling $\hat{\mu}_{it}$. By choosing the same re-sampling scheme for all cross-sectional units, the contemporaneous correlation structure is preserved.

The bootstrap series y^*_{it} and X^*_{it} are generated by first inserting μ^*_{it} in (10) and by then inserting the resulting v^*_{it} in (8) and (9).

The multivariate implementation of the RBB bootstrap is based on an unrestricted VAR(1) for $Z_{it} = [y_{it}, X'_{it}]'$ as follows.

(i) Estimate the first order VAR $Z_{it} = A_{i0} + A_{i1} Z_{it-1} + v_{it}$.

(ii) Compute the centered residuals

$$\tilde{v}_{it} = (Z_{it} - \hat{A}_{i1} Z_{it-1}) - \frac{1}{T-1} \sum_{\tau=2}^{T} (Z_{i\tau} - \hat{A}_{i1} Z_{i\tau-1}).$$

Choose the block-length b and draw $j_0, \dots, j_{k-1}$ from the discrete uniform distribution over the set $\{1, \dots, T-b\}$ with $k = \lfloor \frac{T-1}{b} \rfloor$ and $\lfloor x \rfloor$ denotes the integer part of x. By taking the same realizations j_m for all cross-sections, the cross-sectional correlation is preserved in the bootstrap data.

(iv) Denoting with $m = \lfloor \frac{t-2}{b} \rfloor$ and with $s = t - mb - 1$, the bootstrap data are given by:

$$Z^*_{it} = \begin{cases} Z_{i1} & t = 1 \\ \hat{A}_{i0} + Z^*_{it-1} + \tilde{v}_{ij_m+s} & t = 2, \dots, kb+1 \end{cases} \tag{11}$$

Note again for completeness that for the tests that only allow for an intercept in the test equation $\hat{A}_{i0}$ above is replaced by zero.

Author: Martin Wagner

Title: The Carbon Kuznets Curve: A Cloudy Picture Emitted by Bad Econometrics?

Reihe Ökonomie / Economics Series 197

ISSN: 1605-7996

ISSN: 1605-7996